Universitext

The Special Theory of Relativity

A Mathematical Exposition

狭义相对论的数学表述

A. Das

Springer-Verlag

世界图书出版公司

Anadijiban Das

The Special Theory of Relativity

A Mathematical Exposition

With 27 Illustrations

Springer-Verlag
世界图书出版公司
北京·广州·上海·西安

Anadijiban Das
Department of Mathematics and Statistics
Simon Fraser University
Burnaby, V5A 1S6 British Columbia
Canada

Library of Congress Cataloging-in-Publication Data
Das, Anadijiban.
The special theory of relativity: a mathematical exposition / Anadijiban Das, author.
p. cm.
Includes bibliographical references and index.
ISBN 0-387-94042-1. --ISBN 3-540-94042-1
1. Special relativity (Physics)--Mathematics. 2. Mathematical physics. I. Title.
QC173.65.D38 1993
530.1'1--dc20 93-10256

Printed on acid-free paper.

Reprinted in China by Beijing World Publishing Corporation, 1998

ISBN 0-387-94042-1 Springer-Verlag New York Berlin Heidelberg
ISBN 3-540-94042-1 Springer-Verlag Berlin Heidelberg · New York SPIN 10528157

Preface

The material in this book is presented in a logical sequence rather than a historical sequence. Thus, we feel obligated to sketch briefly the history of the special theory of relativity. The brilliant experiments of Michelson and Morley in 1887 demonstrated the astonishing fact that the speed of light is independent of the state of relative linear motion of the source of light and the observer of the light. This fact necessitates the modification of the usual Galilean transformation (between two relatively moving observers), which tacitly assumes that time and space are absolute.

Fitzgerald in 1889 and Lorentz in 1892 altered the Galilean transformation by introducing a length contraction in the direction of relative motion. This contraction explained the Michelson–Morley experiment, but it was viewed by both Fitzgerald and Lorentz as a mathematical trick only and *not* indicative of the nature of reality. In 1898 Larmor introduced a similar time dilation in an attempt to find the transformations which leave Maxwell's equations invariant. Lorentz also introduced the time dilation independently sometime before 1904. Poincaré in 1905 also discovered the Lorentz transformation and asserted that it was the fundamental invariance group of nature. Einstein in 1905 discovered the Lorentz transformation from physical considerations. Einstein, alone among these mathematical physicists, recognized the philosophical implications of the Lorentz transformation in that it rejected the commonly held notion that space and time were both absolute. He postulated the equivalence of all inertial frames of reference (moving with constant velocities relative to each other) with regard to the formulation of natural laws. Furthermore, he recognized and postulated that the speed of light is the maximum speed of propagation of any physical action. Therefore, the speed of light must be invariant for all inertial observers. Thus the Michelson-Moreley experiment was reconciled with theory. Minkowski, a mathematician, combined both physical postulates of Einstein into one mathematical axiom. This axiom is that "all natural laws must be expressible as tensor field equations on a (flat) absolute space–time manifold." Thus, in that there is no preferred inertial frame for the formulation of natural laws, a universal democracy is postulated to exist among all inertial observers. This axiom is called the Principle of Special Relativity. Many experiments involving atoms and subatomic particles have verified the essential validity of this principle.

In the first chapter we introduce axiomatically the four-dimensional Minkowski vector space. This vector space is endowed with a nondegenerate inner product which is *not* positive definite. Therefore, the concepts of the *norm* (or length) of a four-vector and of the *angle* between two four-vectors have to be *abandoned.* A Lorentz mapping is introduced as an inner product preserving linear mapping of Minkowski vector space into itself.

In Chapter 2 we introduce the flat Minkowski space–time manifold with a proper axiomatic structure. It is proved that the transformation from one Minkowski chart to another must be given by a Poincaré transformation (or an inhomogeneous Lorentz transformation). The *conceptual difference* between a Lorentz transformation of coordinate charts and a Lorentz mapping of the tangent (Minkowski) vector space is clearly displayed. Minkowski tensor fields on the flat space–time are also defined.

In the third chapter, by applications of a particular Lorentz transformation (the "boost"), length contraction, time retardation, and the composition of velocities are explained. The group structure of the set of all Lorentz transformations is demonstrated, and real representations of the Lorentz group are presented. The proper orthochronous subgroup is defined and discussed also.

The fourth chapter defines the spinor space (a two-dimensional complex vector space) and the properties of spinors. Bispinor space (a four-dimensional complex vector space) is also introduced. It is shown that a unimodular mapping of spinor space can induce a proper, orthochronous Lorentz mapping on Minkowski vector space. Furthermore, a unimodular mapping of bispinor space is shown to induce a general Lorentz mapping of Minkowski vector space.

In Chapter 5 prerelativistic mechanics is briefly reviewed. In the setting of prerelativistic mechanics in space and time, $\mathbb{E}_3 \times \mathbb{R}$, the momentum conjugate to the time variable turns out to be the negative of energy! After this, the relativistic mechanics is investigated. The Lagrangian is assumed to be a positive homogeneous function of degree one in the velocity variables (which makes the generalized Hamiltonian identically zero!). Examples from electromagnetic theory and the linearized gravitational theory of Einstein are worked out.

In Chapter 6 the relativistic (classical) field theory is developed. Noether's theorem (essential for the differential conservation laws) is rigorously proved. As examples of special fields, the Klein–Gordon scalar field, the electromagnetic tensor field, nonabelian gauge fields, and the Dirac bispinor field are presented. However, at the present level of treatment, gauge fields are *not* derived as connections in a fibre bundle over the base (Minkowski) manifold. In each chapter, examples and exercises of various degrees of difficulty are provided.

Chapter 7 deals with a research topic, namely, classical fields in the eight-dimensional extended (or covariant) phase space. Historically, Born and Yukawa advocated the extended phase space on the basis of the principle of *reciprocity* (covariance under the canonical transformation $\hat{p} = -q$, $\hat{q} = p$). In recent years, Caianello and others have considered the principle of *maximal proper acceleration* arising out of the extended phase space geometry. We ourselves have done some research on classical fields in the eight-dimensional phase space. We can obtain, in a certain sense, a unified meson field and a unification of fermionic fields. These fields, however, contain *infinitely* many modes or particles.

We have *changed* the usual notation for the Lorentz metric η_{ij} in favor of d_{ij} (since η_{ijkl} is used for the pseudotensor) and $\gamma \equiv (1 - v^2)^{-1/2}$ in favor of $\beta \equiv (1 - v^2)^{-1/2}$ (since γ is used to denote a curve).

This book has grown out of lectures delivered at Jadavpur University (Calcutta), University College of Dublin, Carnegie–Mellon University, and mostly at Simon Fraser University (Canada). The material is intended mainly for students at the fourth and the fifth year university level. We have taken special care to steer a *middle course* between abstruse mathematics and theoretical physics, so that this book can be used for courses in special relativity in both mathematics and physics departments. Furthermore, the material presented here is a suitable prerequisite for further study in either general relativity or relativistic particle theory.

In conclusion, I would like to acknowledge gratefully several people for various reasons. I was fortunate to learn the subject of special relativity from the late Professor S. N. Bose F.R.S. (of Bose–Einstein statistics) in Calcutta University. I also had the privilege for three years of being a research associate of the late Professor J. L. Synge F.R.S. at the Dublin Institute for Advanced Studies. Their influence, direct or indirect, is evident in the presentation of the material (although the errors in the book are solely due to me!). In preparation of the manuscript, I have been helped very much by Dr. Ted Biech, who typed the manuscript and suggested various improvements. Mrs. J. Fabricius typed the difficult Chapter 7. Mrs. E. Carefoot drew the diagrams. Dr. Shounak Das has suggested some literary improvements. I also owe thanks to many of my students for stimulating discussions during lectures.

I thank Dr. S. Kloster for the careful proof reading.

Finally, I thank my wife Mrs. Purabi Das for constant encouragement.

Contents

1
Four-Dimensional Vector Spaces and Linear Mappings

1.1. Minkowski Vector Space $\mathbf{V}_4$

The three-dimensional vectors in Newtonian physics are generalized into four-dimensional vectors in the theory of relativity. This four-dimensional vector space is called the *Minkowski vector space* and is denoted by $\mathbf{V}_4$. This vector space is over the real field $\mathbb{R}$. The mathematical axioms for addition and scalar multiplication of Minkowski vectors are as follows:

A1. $\mathbf{a} + \mathbf{b} \in \mathbf{V}_4$ for all $\mathbf{a}, \mathbf{b} \in \mathbf{V}_4$.

A2. $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$ for all $\mathbf{a}, \mathbf{b} \in \mathbf{V}_4$.

A3. $(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c})$ for all $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbf{V}_4$.

A4. There is $\mathbf{0} \in \mathbf{V}_4$ such that $\mathbf{a} + \mathbf{0} = \mathbf{a}$ for all $\mathbf{a} \in \mathbf{V}_4$.

A5. For all $\mathbf{a} \in \mathbf{V}_4$ there is $-\mathbf{a} \in \mathbf{V}_4$ so that $(-\mathbf{a}) + \mathbf{a} = 0$. (1.1.1)

M1. $\alpha\mathbf{a} \in \mathbf{V}_4$ for all $\alpha \in \mathbb{R}$, for all $\mathbf{a} \in \mathbf{V}_4$.

M2. $\alpha(\beta\mathbf{a}) = (\alpha\beta)\mathbf{a}$ for all $\alpha, \beta \in \mathbb{R}$, for all $\mathbf{a} \in \mathbf{V}_4$.

M3. $1\mathbf{a} = \mathbf{a}$ for all $\mathbf{a} \in \mathbf{V}_4$.

M4. $\alpha(\mathbf{a} + \mathbf{b}) = \alpha\mathbf{a} + \alpha\mathbf{b}$ for all $\alpha \in \mathbb{R}$, for all $\mathbf{a}, \mathbf{b} \in \mathbf{V}_4$.

M5. $(\alpha + \beta)\mathbf{a} = \alpha\mathbf{a} + \beta\mathbf{a}$ for all $\alpha, \beta \in \mathbb{R}$, for all $\mathbf{a} \in \mathbf{V}_4$.

We shall also assume the existence of an inner product for $\mathbf{V}_4$ satisfying the following axioms:

I1. $\mathbf{a} \cdot \mathbf{b} \in \mathbb{R}$ for all $\mathbf{a}, \mathbf{b} \in \mathbf{V}_4$.

I2. $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$ for all $\mathbf{a}, \mathbf{b} \in \mathbf{V}_4$.

I3. $(\lambda\mathbf{a} + \mu\mathbf{b}) \cdot \mathbf{c} = \lambda(\mathbf{a} \cdot \mathbf{c}) + \mu(\mathbf{b} \cdot \mathbf{c})$ for all $\lambda, \mu \in \mathbb{R}$, for all $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbf{V}_4$. (1.1.2)

I4. $\mathbf{a} \cdot \mathbf{x} = 0$ for all $\mathbf{x} \in \mathbf{V}_4$ if and only if $\mathbf{a} = \mathbf{0}$.

The axiom I4′ is called the *axiom of nondegeneracy*. It is a weaker axiom than

I5. $\mathbf{a}\cdot\mathbf{a} \geq 0$ for all $\mathbf{a} \in \mathbf{V}_4$, and $\mathbf{a}\cdot\mathbf{a} = 0$ if and only if $\mathbf{a} = 0$. (1.1.3)

For a *positive definite* inner product axiom I5 replaces I4. In addition to these axioms we impose the *axiom of dimensionality* on Minkowski vector space:

D1. $\dim \mathbf{V}_4 = 4$.

Let $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$ be a basis set for $\mathbf{V}_4$. The *metric tensor components* relative to this basis are defined by

$$g_{ij} \equiv \mathbf{e}_i \cdot \mathbf{e}_j \quad \text{for } i, j \in \{1,2,3,4\}. \tag{1.1.4}$$

From axiom I2 it follows that $g_{ji} = g_{ij}$ for all i, $j \in \{1,2,3,4\}$. The four-dimensional unit matrix is denoted by $\mathbf{I} \equiv [\delta_{ij}]$. The eigenvalues of the matrix $[g_{ij}]$ are the roots of the characteristic equation

$$\det[g_{ij} - \lambda\delta_{ij}] = 0. \tag{1.1.5}$$

Since the matrix g_{ij} is symmetric, the roots of (1.1.5) are all real. By the axiom of nondegeneracy I4 it follows that all the eigenvalues of g_{ij} are nonzero. The signs of the eigenvalues of g_{ij} are determined by the *axiom of Lorentz signature*:

S1. $\lambda_1 > 0, \lambda_2 > 0, \lambda_3 > 0, \lambda_4 < 0$.

The vector space obeying the sixteen axioms A1–A5, M1–M5, I1–I4, D1, and S1 is called *Minkowski vector space* and is denoted by $\mathbf{V}_4$.

In $\mathbf{V}_4$, the two vectors **a**, **b** are defined to be *Minkowski orthogonal* (or *M-orthogonal*) provided

$$\mathbf{a}\cdot\mathbf{b} = 0. \tag{1.1.6}$$

Theorem (1.1.1): *There exists an M-orthonormal basis* $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$ *for* $\mathbf{V}_4$ *such that*

$$g_{ij} = \mathbf{e}_i \cdot \mathbf{e}_j = d_{ij}, \tag{1.1.7}$$

where

$$D = [d_{ij}] \equiv \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}.$$

The proof is rather involved and is omitted. The metric d_{ij} in (1.1.7) is called the *Lorentz metric*. The signature of d_{ij} is defined to be the trace of $[d_{ij}]$. We shall use a choice of $[d_{ij}]$ so that the signature is equal to 2. Note that some authors use the signature -2.

Now we shall explain the *Einstein summation convention*. In a mathematical expression, wherever two repeated Roman indices are present, the sum over the repeated index is implied. For example, we write

$$u^k v_k \equiv \sum_{k=1}^{4} u^k v_k = \sum_{l=1}^{4} u^l v_l = u^l v_l,$$

$$g_{ij} u^i v^j \equiv \sum_{i=1}^{4} \sum_{j=1}^{4} g_{ij} u^i v^j = \sum_{k=1}^{4} \sum_{l=1}^{4} g_{kl} u^k v^l = g_{kl} u^k v^l.$$

The summation indices are called dummy indices, since they can be replaced by other indices over the same range. In the summation convention, *never* use dummy indices that repeat more than twice. This is necessary in order to avoid wrong answers; for example,

$$u^k v_k u^k v_k \equiv \sum_{k=1}^{4} u^k v_k u^k v_k \neq \sum_{k=1}^{4} \sum_{l=1}^{4} u^k v_k u^l v_l = u^k v_k u^l v_l = (u^k v_k)^2.$$

Let $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$ be an M-orthonormal basis (or *tetrad*) for $\mathbf{V}_4$. For any vector $\mathbf{u} \in \mathbf{V}_4$, there exists a linear combination

$$\mathbf{u} = \sum_{i=1}^{4} u^i \mathbf{e}_i = u^i \mathbf{e}_i. \tag{1.1.8}$$

The unique numbers or scalars u^i are called the *Minkowski components* of the vector $\mathbf{u}$ relative to the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$.

Theorem (1.1.2): *In terms of the Minkowski components, the inner product between vectors* $\mathbf{u}$, $\mathbf{v}$ *is given by*

$$\mathbf{u} \cdot \mathbf{v} = d_{ij} u^i v^j. \tag{1.1.9}$$

Proof: Choose an M-orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$ such that

$$\mathbf{u} = u^i \mathbf{e}_i, \qquad \mathbf{v} = v^i \mathbf{e}_i.$$

By the axioms in (1.1.2) and (1.1.8) we have

$$\mathbf{u} \cdot \mathbf{v} = (u^i \mathbf{e}_i) \cdot (v^j \mathbf{e}_j) = d_{ij} u^i v^j. \quad \blacksquare$$

Note that from (1.1.9)

$$\mathbf{u} \cdot \mathbf{u} = d_{ij} u^i u^j = (u^1)^2 + (u^2)^2 + (u^3)^2 - (u^4)^2. \tag{1.1.10}$$

The above expression is *not* positive definite. Thus the concept of the *length* (or norm) of a vector in $\mathbf{V}_4$ *is abandoned*. Furthermore, if we define $\cos(\mathbf{u}, \mathbf{v}) \equiv (\mathbf{u} \cdot \mathbf{v})/\sqrt{(\mathbf{u} \cdot \mathbf{u})(\mathbf{v} \cdot \mathbf{v})}$, then we are led to contradictions. For example, if we choose $\mathbf{v}_n = \mathbf{e}_1 + [(n-1)/n]\mathbf{e}_4$ for $n \in \mathbb{Z}^+$ and let $\mathbf{u} = \mathbf{e}_1$, then $\cos(\mathbf{e}_1, \mathbf{v}_n) \equiv n/\sqrt{(2n-1)}$. Therefore, $1 \leq \cos(\mathbf{u}, \mathbf{v}_n)$ and $\lim_{n\to\infty} \cos(\mathbf{u}, \mathbf{v}_n) \to \infty$, which is absurd. That is why the concept of an *angle* between two vectors $\mathbf{u}, \mathbf{v} \in \mathbf{V}_4$ *is abandoned* as well. However, for a spatial vector subspace $\mathbf{V}_3 \equiv \{\mathbf{v} \in \mathbf{V}_4 : v^4 = 0\}$, the usual concept of the length and angle can be restored.

Since the expression (1.1.10) for $\mathbf{u} \cdot \mathbf{u}$ is indefinite, we can define three kinds of vectors in $\mathbf{V}_4$:

(i) a vector $\mathbf{u} \in \mathbf{V}_4$ that satisfies $\mathbf{u} \cdot \mathbf{u} > 0$ is called a *spacelike* vector;
(ii) a vector $\mathbf{u} \in \mathbf{V}_4$ for which $\mathbf{u} \cdot \mathbf{u} < 0$ is called a *timelike* vector;
(iii) a vector $\mathbf{u} \in \mathbf{V}_4$ for which $\mathbf{u} \cdot \mathbf{u} = 0$ is called a *null* vector.

Example: Let $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$ be an M-orthonormal basis for $\mathbf{V}_4$. By (1.1.9), $\mathbf{e}_1 \cdot \mathbf{e}_1 = d_{11} = 1$. Thus $\mathbf{e}_1$ is a spacelike vector. Similarly $\mathbf{e}_2$, $\mathbf{e}_3$ are spacelike vectors. But $\mathbf{e}_4 \cdot \mathbf{e}_4 = d_{44} = -1$, so $\mathbf{e}_4$ is a timelike vector. Set $\mathbf{u} = \mathbf{e}_1 + \mathbf{e}_4$ and observe that $\mathbf{u} \cdot \mathbf{u} = 0$, so we see that $\mathbf{u}$ is a null vector. □

The *separation number* is a generalization of the concept of length and is denoted by $\sigma(\mathbf{u})$. It is defined by

$$\sigma(\mathbf{u}) \equiv \sqrt{|\mathbf{u} \cdot \mathbf{u}|} \geq 0. \tag{1.1.11}$$

Thus for either timelike or spacelike vectors $\mathbf{u} \in \mathbf{V}_4$ we have $\sigma(\mathbf{u}) > 0$. But for a null vector $\mathbf{n}$ we have $\sigma(\mathbf{n}) = 0$. For an example choose $\mathbf{u} = (\mathbf{e}_1 - 2\mathbf{e}_4)/2$. Then $\sigma(\mathbf{u}) = \sqrt{|-3/4|} = \sqrt{3}/2$.

A vector $\mathbf{e}$ in $\mathbf{V}_4$ is called a *unit vector* if $\sigma(\mathbf{e}) = 1$. Subsequently we shall use only M-orthonormal bases for $\mathbf{V}_4$ unless mentioned otherwise. A spatial vector subspace relative to an M-orthonormal basis is defined as

$$\mathbf{V}_3 \equiv \{\mathbf{v} \in \mathbf{V}_4 \colon v^4 = 0\}. \tag{1.1.12}$$

Small Greek indices will take values in the set $\{1, 2, 3\}$, and *small Roman indices* will take values in the set $\{1, 2, 3, 4\}$. The appropriate summation convention will apply to each type of index.

Theorem (1.1.3) (Schwarz Inequality): *For any two vectors* $\mathbf{u}$, $\mathbf{v}$ *in* $\mathbf{V}_3$ *the following inequality holds*:

$$|u^\alpha v^\alpha| \leq \sqrt{u^\alpha u^\alpha v^\beta v^\beta}. \tag{1.1.13}$$

Equality holds if and only if $u^\alpha = \lambda v^\alpha$ *for some* $\lambda \in \mathbb{R}$.

Proof: Suppose that $u^\alpha \equiv 0$. Then (1.1.13) holds trivially. Now suppose $u^\alpha u^\alpha > 0$. Then for any $\lambda \in \mathbb{R}$ we have

$$(\lambda u^\alpha + v^\alpha)(\lambda u^\alpha + v^\alpha) = \lambda^2 u^\alpha u^\alpha + v^\alpha v^\alpha + 2\lambda u^\alpha v^\alpha \geq 0.$$

Setting the value $\lambda = -(u^\alpha v^\alpha)/(u^\gamma u^\gamma)$, we obtain

$$[-(u^\alpha v^\alpha)^2 + (u^\alpha u^\alpha)(v^\beta v^\beta)]/(u^\gamma u^\gamma) \geq 0.$$

From above (1.1.13) follows. The case of equality mentioned in the theorem is left as an exercise. ■

The M-orthogonality between two vectors in $\mathbf{V}_4$ is not always intuitively natural. We shall derive a few theorems on that topic now.

Theorem (1.1.4): *No two timelike vectors in* $\mathbf{V}_4$ *can be M-orthogonal.*

Proof: Let **u**, **v** be two timelike vectors. Thus we have

$$\mathbf{u}\cdot\mathbf{u} = d_{ij}u^i u^j = u^\alpha u^\alpha - (u^4)^2 < 0,$$

$$\mathbf{v}\cdot\mathbf{v} = d_{ij}v^i v^j = v^\alpha v^\alpha - (v^4)^2 < 0.$$

Combining these two inequalities we have

$$\sqrt{u^\alpha u^\alpha v^\beta v^\beta} < |u^4 v^4|.$$

By the Schwarz inequality (1.1.13) we obtain

$$|u^\alpha v^\alpha| < |u^4 v^4|. \tag{1.1.14}$$

Suppose, contrary to the conclusion of the theorem, that $\mathbf{u}\cdot\mathbf{v} = 0$. Then $|u^\alpha v^\alpha| = |u^4 v^4|$. This last equality contradicts (1.1.14). ■

Corollary (1.1.1): *For two timelike vectors* **u**, **v** *such that* $u^4 > 0$, $v^4 > 0$, *we have* $\mathbf{u}\cdot\mathbf{v} < 0$.

Proof: $u^\alpha v^\alpha \leq |u^\alpha v^\alpha| \leq \sqrt{u^\alpha u^\alpha v^\beta v^\beta} < |u^4 v^4| = u^4 v^4$, so $\mathbf{u}\cdot\mathbf{v} < 0$. ■

Theorem (1.1.5) (Synge): *Let* $\hat{\mathbf{t}}$, $\mathbf{t}$ *be two timelike, future-pointing unit vectors (past-pointing may replace future-pointing). Then* $-\infty < \hat{\mathbf{t}}\cdot\mathbf{t} \leq -1$.

Proof: For definiteness assume that $\hat{\mathbf{t}}$, $\mathbf{t}$ are future-pointing timelike vectors. Thus

$$\sigma(\mathbf{t}) = 1, \qquad t^\alpha t^\alpha - (t^4)^2 = -1, \qquad t^4 > 0,$$

$$\sigma(\hat{\mathbf{t}}) = 1, \qquad \hat{t}^\beta \hat{t}^\beta - (\hat{t}^4)^2 = -1, \qquad \hat{t}^4 > 0.$$

We want to solve the above equations and inequalities. For that purpose consider two three-dimensional unit spatial vectors:

$$a^\alpha a^\alpha = 1, \qquad \hat{a}^\alpha \hat{a}^\alpha = 1.$$

In spherical polar coordinates we can write

$$a^1 = \sin\theta\cos\phi, \qquad \hat{a}^1 = \sin\hat{\theta}\cos\hat{\phi},$$

$$a^2 = \sin\theta\sin\phi, \qquad \hat{a}^2 = \sin\hat{\theta}\sin\hat{\phi},$$

$$a^3 = \cos\theta, \qquad \hat{a}^3 = \cos\hat{\theta},$$

where $0 \leq \theta \leq \pi$, $0 \leq \hat{\theta} \leq \pi$, $-\pi \leq \phi < \pi$, and $-\pi \leq \hat{\phi} < \pi$. Thus we have $\cos\psi = a^\alpha \hat{a}^\alpha = \cos\theta\cos\hat{\theta} + \sin\theta\sin\hat{\theta}\cos(\phi - \hat{\phi})$, where ψ is the angle between the two vectors a^ρ, $\hat{a}^\rho$ and $0 \leq \psi \leq \pi$. We can always express **t** and $\hat{\mathbf{t}}$ in the form

$$t^\alpha = (\sinh\chi)a^\alpha, \qquad \hat{t}^\alpha = (\sinh\hat{\chi})\hat{a}^\alpha,$$

$$t^4 = \cosh\chi > 0, \qquad \hat{t}^4 = \cosh\hat{\chi} > 0,$$

where $\chi, \hat{\chi} \in \mathbb{R}$. Therefore,

$$\begin{aligned}-\mathbf{t}\cdot\hat{\mathbf{t}} &= \cosh\chi\cosh\hat{\chi}[\cos^2(\psi/2) + \sin^2(\psi/2)] \\ &\quad - \sinh\chi\sinh\hat{\chi}[\cos^2(\psi/2) - \sin^2(\psi/2)] \\ &= \cosh(\chi - \hat{\chi})\cos^2(\psi/2) + \cosh(\chi + \hat{\chi})\sin^2(\psi/2) > 0.\end{aligned}$$

Making another transformation $x = \sin(\psi/2)$ so that $x \in [0, 1]$, we get $y \equiv -\mathbf{t}\cdot\hat{\mathbf{t}} = [\cosh(\chi + \hat{\chi}) - \cosh(\chi - \hat{\chi})]x^2 + \cosh(\chi - \hat{\chi}) \equiv f(x)$. Thus we can graph the function $f(x)$ over $x \in [0, 1]$. For the case $\cosh(\chi + \hat{\chi}) - \cosh(\chi - \hat{\chi}) > 0$, this graph is a portion of a parabola with $f'(x) \geq 0$ and $\cosh(\chi - \hat{\chi}) \leq y \leq \cosh(\chi + \hat{\chi})$. For the case $\cosh(\chi + \hat{\chi}) - \cosh(\chi - \hat{\chi}) < 0$, the graph is a portion of a parabola with $f'(x) \leq 0$ and $\cosh(\chi + \hat{\chi}) \leq y \leq \cosh(\chi - \hat{\chi})$. In the case $\cosh(\chi + \hat{\chi}) - \cosh(\chi - \hat{\chi}) = 0$, the graph is a portion of a straight line with $f'(x) \equiv 0$ and $y = \cosh(\chi + \hat{\chi}) = \cosh(\chi - \hat{\chi})$. In all three cases $\min\{\cosh(\chi + \hat{\chi}), \cosh(\chi - \hat{\chi})\} \leq y \leq \max\{\cosh(\chi + \hat{\chi}), \cosh(\chi - \hat{\chi})\}$. Since the function $\cosh(x)$ has the lower bound 1 and no upper bound, we see that $1 \leq y < \infty$, so $-\infty < \mathbf{t}\cdot\hat{\mathbf{t}} \leq -1$. ■

Corollary (1.1.2): *Let* $\mathbf{t}, \hat{\mathbf{t}}$ *be two timelike, future-pointing unit vectors such that* $\mathbf{t}\cdot\hat{\mathbf{t}} = -1$. *Then* $\mathbf{t} = \hat{\mathbf{t}}$.

Proof: From the proof of Theorem (1.1.5), it is clear that, for the case $\cosh(\chi + \hat{\chi}) > \cosh(\chi - \hat{\chi})$, the minimum value $y = 1$ is at $x = 0$ and the minimum value is $\cosh(\chi - \hat{\chi}) = 1$. Thus $\psi = 0$ and $\chi = \hat{\chi}$, so $\theta = \hat{\theta}$, $\phi = \hat{\phi}$. Thus $t^\alpha = \hat{t}^\alpha$, $t^4 = \hat{t}^4$; hence, $\mathbf{t} = \hat{\mathbf{t}}$. In the second case $\cosh(\chi + \hat{\chi}) < \cosh(\chi - \hat{\chi})$, and the minimum value is $\cosh(\chi + \hat{\chi}) = 1$ at $x = 1$. Therefore, $\psi = \pi$ and $\chi = -\hat{\chi}$ and so $\hat{\theta} = \pi - \theta$, $\hat{\phi} = \phi + \pi$. Thus $t^\alpha = \hat{t}^\alpha$, $t^4 = \hat{t}^4$, so $\mathbf{t} = \hat{\mathbf{t}}$. In the case $\cosh(\chi + \hat{\chi}) = \cosh(\chi - \hat{\chi})$, the minimum value is 1, which is attained by all $x \in [0, 1]$. Then $\cosh(\chi + \hat{\chi}) = \cosh(\chi - \hat{\chi}) = 1$; hence, $\chi = \hat{\chi} = 0$. So $t^\alpha = \hat{t}^\alpha = 0$, $t^4 = \hat{t}^4$, so $\mathbf{t} = \hat{\mathbf{t}}$. ■

Theorem (1.1.6): *A timelike vector cannot be M-orthogonal to a nonzero null vector.*

Proof: Suppose that $\mathbf{t}$ is a timelike vector and $\mathbf{n}$ is a nonzero null vector. In terms of their Minkowski components $t^\alpha t^\alpha < (t^4)^2$ and $n^\alpha n^\alpha = (n^4)^2$ with $|t^4| > 0$, $|n^4| > 0$. Combining these expressions and using the Schwarz inequality (1.1.13) we have

$$(t^\alpha n^\alpha)^2 \leq t^\alpha t^\alpha n^\beta n^\beta < (t^4 n^4)^2. \tag{1.1.15}$$

Suppose that $\mathbf{t}\cdot\mathbf{n} = 0$, or $(t^\alpha n^\alpha)^2 = (t^4 n^4)^2$. This contradicts (1.1.15). ■

Now we shall prove a *very counterintuitive theorem.*

Theorem (1.1.7): *Two nonzero null vectors are M-orthogonal if and only if they are scalar multiples of each other.*

Proof: (i) Assume that two null vectors $\mathbf{m}$, $\mathbf{n}$ are such that $\mathbf{m} = \lambda\mathbf{n}$ for some $\lambda \in \mathbb{R}$. Then $\mathbf{m}\cdot\mathbf{n} = \lambda(\mathbf{n}\cdot\mathbf{n}) = 0$.

(ii) Suppose that two nonzero null vectors $\mathbf{m}$, $\mathbf{n}$ are M-orthogonal. Then

$$\begin{aligned} \mathbf{m}\cdot\mathbf{m} &= m^\alpha m^\alpha - (m^4)^2 = 0, \\ \mathbf{n}\cdot\mathbf{n} &= n^\alpha n^\alpha - (n^4)^2 = 0, \\ \mathbf{m}\cdot\mathbf{n} &= m^\alpha n^\alpha - m^4 n^4 = 0, \qquad m^4 \neq 0, n^4 \neq 0. \end{aligned} \tag{1.1.16}$$

From the above expressions we obtain

$$(m^\alpha n^\alpha)^2 = (m^4 n^4)^2 = m^\alpha m^\alpha n^\beta n^\beta, \qquad |m^\alpha n^\alpha| = \sqrt{m^\alpha m^\alpha n^\beta n^\beta}. \tag{1.1.17}$$

The above equation is the case of equality in the Schwartz inequality (1.1.13). Therefore, $m^\alpha = \lambda n^\alpha$ for some scalar $\lambda \neq 0$. Since $n^4 \neq 0$, we have $m^4 = m^\alpha n^\alpha/n^4 = \lambda n^\alpha n^\alpha/n^4 = \lambda n^4$. Thus $\mathbf{m} = \lambda\mathbf{n}$. ■

It is hard to plot the Minkowski vectors, since the concepts of length of a vector and angle between two vectors do not exist. However, the *parallelogram law* of vector addition still holds. It is worthwhile to draw Minkowski vectors to gain some geometrical insight. We have to plot these vectors on a piece of paper, which is part of a Euclidean plane. Let us plot M-orthonormal vectors $\mathbf{e}_1, \mathbf{e}_4$ such that $\mathbf{e}_1\cdot\mathbf{e}_4 = 0$ and $\mathbf{e}_1\cdot\mathbf{e}_1 = -\mathbf{e}_4\cdot\mathbf{e}_4 = 1$. It is quite natural to plot these two vectors as $\mathbf{i}$ and $\mathbf{j}$ of the usual two-dimensional Cartesian basis vectors; see Figure 1. As we have drawn, the Euclidean lengths $\|\mathbf{e}_1\| = \|\mathbf{e}_4\| = 1$ and $\mathbf{e}_1, \mathbf{e}_4$ are Euclidean orthogonal. However, the vectors $\mathbf{e}_1 + \mathbf{e}_4$, $-\mathbf{e}_1 + \mathbf{e}_4$ have Euclidean lengths $\|\mathbf{e}_1 + \mathbf{e}_4\| = \|-\mathbf{e}_1 + \mathbf{e}_4\| = \sqrt{2}$, and $\mathbf{e}_1 + \mathbf{e}_4$, $-\mathbf{e}_1 + \mathbf{e}_4$ are Euclidean orthogonal. But $\sigma(\mathbf{e}_1 + \mathbf{e}_4) = 0$, $\sigma(-\mathbf{e}_1 + \mathbf{e}_4) = 0$, and $(\mathbf{e}_1 + \mathbf{e}_4)\cdot(-\mathbf{e}_1 + \mathbf{e}_4) = -2 \neq 0$. So we have to *use caution* in order to interpret any plot of Minkowski vectors.

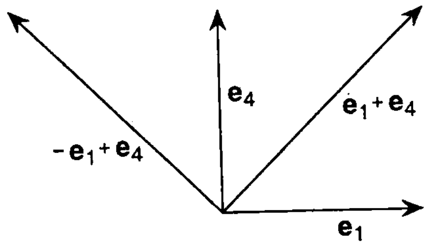

FIGURE 1. Minkowski vectors $\mathbf{e}_1$, $\mathbf{e}_4$, $\mathbf{e}_1 + \mathbf{e}_4$, $-\mathbf{e}_1 + \mathbf{e}_4$.

EXERCISES 1.1

1. Let $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$ be a Minkowski basis for $\mathbf{V}_4$. Let another basis be $\{\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}\}$ where

$$\mathbf{a} = \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3 - \mathbf{e}_4, \qquad \mathbf{b} = \mathbf{e}_2 - \mathbf{e}_3 + \sqrt{2}\mathbf{e}_4,$$
$$\mathbf{c} = \mathbf{e}_3 - \mathbf{e}_4, \qquad \mathbf{d} = \mathbf{e}_4.$$

(i) Determine which of these are spacelike, timelike, or null vectors.
(ii) Obtain the separation numbers $\sigma(\mathbf{a})$, $\sigma(\mathbf{b})$, $\sigma(\mathbf{c})$, $\sigma(\mathbf{d})$.
(iii) Determine whether or not $\{\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}\}$ is a basis for $\mathbf{V}_4$.

2. Determine which of the following subsets of $\mathbf{V}_4$ is a vector subspace.
(i) The union of the subset of all spacelike vectors and $\{\mathbf{0}\}$.
(ii) The union of the subset of all timelike vectors and $\{\mathbf{0}\}$.
(iii) The subset of all null vectors.

3. Prove that the nondegeneracy axiom of the inner product in (1.1.2) implies that $\det[g_{ij}] \neq 0$.

4. Prove that for any two vectors $\mathbf{x}$, $\mathbf{y}$ in $\mathbf{V}_3$ that

$$|x^\alpha y^\alpha| = \sqrt{x^\alpha x^\alpha y^\beta y^\beta}$$

holds if and only if $\mathbf{x} = \lambda \mathbf{y}$ for some $\lambda \in \mathbb{R}$.

5. Let $\mathbf{s}$, $\mathbf{t}$ be two timelike future-pointing vectors in $\mathbf{V}_4$. Prove that $\sigma(\mathbf{s})\sigma(\mathbf{t}) \leq |\mathbf{s} \cdot \mathbf{t}|$. (This is called the *Reversed Schwarz Inequality* for timelike future-pointing vectors.) Does it hold for other cases, i.e., past-pointing or mixed orientation?

1.2. Lorentz Mappings of $\mathbf{V}_4$

A *linear mapping* $\mathbf{L}: \mathbf{V}_4 \to \mathbf{V}_4$ is defined to be such that

$$\mathbf{L}(\lambda \mathbf{a} + \mu \mathbf{b}) = \lambda \mathbf{L}(\mathbf{a}) + \mu \mathbf{L}(\mathbf{b}) \tag{1.2.1}$$

for all $\lambda, \mu \in \mathbb{R}$ and all $\mathbf{a}, \mathbf{b} \in \mathbf{V}_4$.

Suppose that $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$ is a basis for $\mathbf{V}_4$ that is not necessarily M-orthonormal. Since $\mathbf{L}(\mathbf{e}_j) \in \mathbf{V}_4$, we must have

$$\hat{\mathbf{e}}_j \equiv \mathbf{L}(\mathbf{e}_j) = l^i{}_j \mathbf{e}_i, \tag{1.2.2}$$

for some suitable scalars $l^i{}_j$, $i, j \in \{1, 2, 3, 4\}$, and the summation convention applies. The 4×4 matrix $L \equiv [l^i{}_j]$, where i stands for the row index and j for the column index, is called the *matrix representation* of the mapping $\mathbf{L}$ relative to the bases of $\mathbf{V}_4$, $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$, and $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3, \hat{\mathbf{e}}_4\}$.

Theorem (1.2.1): *Let* $\hat{\mathbf{e}}_i \equiv \mathbf{L}(\mathbf{e}_i)$. *The set of vectors* $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3, \hat{\mathbf{e}}_4\}$ *is also a basis for* $\mathbf{V}_4$ *if and only if* $\det[l^i{}_j] \neq 0$.

The proof is left to the reader.

A linear mapping $\mathbf{L}$ with $\det[l^i{}_j] \neq 0$ is called *invertible*.

Example: Let us consider a linear mapping $\mathbf{L}$ such that

$$\hat{\mathbf{e}}_1 = \mathbf{L}(\mathbf{e}_1) \equiv \sin(\pi/4)\mathbf{e}_1 + \cos(\pi/4)\mathbf{e}_2 = l^i{}_1 \mathbf{e}_i,$$

$$\hat{\mathbf{e}}_2 = \mathbf{L}(\mathbf{e}_2) \equiv -\cos(\pi/4)\mathbf{e}_1 + \sin(\pi/4)\mathbf{e}_2 = l^i{}_2 \mathbf{e}_i,$$

$$\hat{\mathbf{e}}_3 = \mathbf{L}(\mathbf{e}_3) \equiv \mathbf{e}_3 = l^i{}_3 \mathbf{e}_i,$$

$$\hat{\mathbf{e}}_4 = \mathbf{L}(\mathbf{e}_4) \equiv \mathbf{e}_4 = l^i{}_4 \mathbf{e}_i.$$

Therefore, the 4×4 matrix representation is

$$L = [l^i{}_j] = \begin{bmatrix} \sin(\pi/4) & -\cos(\pi/4) & 0 & 0 \\ \cos(\pi/4) & \sin(\pi/4) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

with $\det[l^i{}_j] = 1$. The linear mapping $\mathbf{L}$ is invertible. The basis $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3, \hat{\mathbf{e}}_4\}$ is M-orthonormal whenever $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$ is M-orthonormal. □

Now we shall define the *Kronecker delta*:

$$\delta^i{}_j \equiv \begin{cases} 1 & \text{for } i = j, \\ 0 & \text{for } i \neq j. \end{cases} \tag{1.2.3}$$

These numbers are the entries of the 4×4 identity matrix I relative to the standard basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$. Similarly $\delta^{\alpha\beta}$ are the entries of the 3×3 identity matrix with respect to the standard basis $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$. We shall work out some examples involving the Kronecker delta.

Example: Consider the sum

$$\begin{aligned} \delta^1{}_j u^j &= \delta^1{}_1 u^1 + \delta^1{}_2 u^2 + \delta^1{}_3 u^3 + \delta^1{}_4 u^4 \\ &= 1u^1 + 0u^2 + 0u^3 + 0u^4 = u^1. \end{aligned}$$

Similarly

$$\delta^i{}_j u^j = u^i, \tag{1.2.4}$$

$$\delta^i{}_j \delta^j{}_k = \delta^i{}_k. \quad \square \tag{1.2.5}$$

For an invertible linear mapping $\mathbf{L}$, we denote the inverse mapping by $\mathbf{A}$, so that the corresponding matrices satisfy

$$A = L^{-1}, \qquad AL = LA = I, \qquad a^i{}_j l^j{}_k = l^i{}_j a^j{}_k = \delta^i{}_k. \tag{1.2.6}$$

Example:

$$[l^i{}_j] = \begin{bmatrix} \sin(\pi/4) & \cos(\pi/4) & 0 & 0 \\ -\cos(\pi/4) & \sin(\pi/4) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$[a^i{}_j] = \begin{bmatrix} \sin(\pi/4) & -\cos(\pi/4) & 0 & 0 \\ \cos(\pi/4) & \sin(\pi/4) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad \square$$